

## FOCUS AND MULTIPLE CPs IN ENGLISH AND BULGARIAN<sup>1</sup>

Elly van Gelderen and Lily Grozeva

University of Groningen and Bulgarian Academy of Sciences

### 1. Introduction

In English and in many other languages, as is well-known, focus does not occur on preposed elements. Focussed phrases introduce new material and the ban on preposing may be due to wanting to link the initial part of a sentence to previously discussed, i.e. old, material. Topicalized items typically refer to information already present in the discourse. Cinque (1993) formalizes this by arguing that the most prominent accent falls on the most embedded constituent in a sentence. There are, however, languages that have focus positions and some of these are at the beginning of a sentence. In Bulgarian, pre-complementizer focus occurs and Hungarian (Horvath 1986) and Chadic (Tuller 1992) provide instances of pre-V and pre-I focus.

In this paper, we examine the structural representation of focus in terms of multiple CP structures. We will argue that structural focus must be checked (cf. Chomsky 1992; 1994) and that the Bulgarian complementizer *li* may have focus features. This is not true in English where C has no structural focus<sup>2</sup>. Thus, structural focus resides in C in Bulgarian and Spec CP 'houses' focus features rather than that a position adjoined to CP does. These pre-C positions are independently needed in the case of multiple fronted *wh*-elements.

The outline is as follows. In section 2, we present the data from English and in section 3, those from Bulgarian as well as our analysis and some problems. In section 4, the focus characteristics of Bulgarian *nali/dali* are discussed and we conclude by saying that there are two parametric differences between Bulgarian and English. The first is that focus features in Bulgarian reside in C but that this is not true in English and the second is that Bulgarian and English differ in the number of functional projections they have available: Bulgarian generates multiple CPs freely whereas English (generally) has one (cf. Authier 1992 and van Gelderen 1993 for differences). We follow Horvath (1986) and Cinque (1993) for English.

### 2. English

Rochemont & Culicover (1990: 18) argue that "all and only the information requested in the [wh]question is focused in the response". Thus, a sentence like (1), where the capitalized words have a focus-related accent, can be the reply to (2). In this paper, we will only be concerned with structural and not intonational focus as in (1). Rochemont & Culicover are

---

<sup>1</sup> We would like to thank Kwangho Lee for his comments.

<sup>2</sup> If one argues that *who* in English and other languages has focus, the preposing may be caused by independent factors such as questioning.

concerned with focus but their criterion on which elements are focussed can be used to separate focus and topic. For instance, (3) is not an appropriate reply to (2), because the fronted position is a typical topic position and not a focus one:

1. I gave FLOWERS to Josie.
2. What did you give to Josie?
3. %FLOWERS, I gave to Josie.

The same is true in Yes/No questions as in (4). Answer (5) is well-formed but not answer (6):

4. Did John give flowers to Andrea?
5. No, John gave flowers to Josie.
6. %No, to Josie, John gave flowers.

A second diagnostic to separate topic and focus is that focussing particles such as *only* and *even* may appear in front of focussed elements, but not preceding topics, as (7) and (8) show:

7. I gave only flowers to Josie.
8. ?Only flowers, I gave to Josie.

Thus the answer to a question is typically new information which can be regarded as the focus of the answer whereas a proposed element cannot. This is different in Bulgarian where the proposed element closest to the left of *li* is focussed. We will therefore argue that structural focus in Bulgarian comes about through the checking of the focus features of C in Spec CP.

In English, focus on NPs is not related to CP or to preposing as (1), (3), (5) and (6) show. I will assume Cinque (1993) who argues that the most deeply embedded element is focussed. For English, this would mean the rightmost. Thus, English focus constructions do not show any evidence for the intricate multiple embedding found in Bulgarian. We might therefore say that Universal Grammar (a) makes available more functional categories in Bulgarian than in English, and (b) that focus features need not be tied to a Focus Phrase but that C can be a host(ess) (cf. also similar conclusions reached in van Gelderen 1993 for different constructions).

### 3. Bulgarian *li*<sup>3</sup>

In Bulgarian, the word order is very free and a Subject, Object and Verb can be combined in six ways, even though SVO is the least marked. Yet, there are some word order rules. When a Verb occurs before *li*, there is no focus but the sentence is a question. When *li* is immediately preceded by an NP, this NP is the focus. In this section, we argue that there is

---

<sup>3</sup> There has recently been a lot of discussion about *li*-lowering (e.g. Rivero 1993). We focus on the structure to the left of C and hence, we need not enter into the *li*-lowering debate.

a special (structural) focus in C, responsible for this.

### 3.1 *Preverbal li*

As mentioned, the following focussing strategies occur (cf. also Rudin 1993/4):

9. Gleda li Marija filma?  
'watch Q. Marija film-the', i.e. Is Marija watching the film?
10. Marija gleda li filma?  
'As for Marija, is she watching the film'?
11. Filma gleda li Marija?  
'As for the film, is Marija watching it'?
12. Marija li gleda filma?  
'Is it Marija who is watching the film'?
13. Filma li gleda Marija?  
'Is it the film that Marija is watching'?
14. Marija filma li gleda?  
'As for Marija, is it the film that she is watching'?
15. Filma Marija li gleda?  
'As for the film, is it Marija who is watching it'?

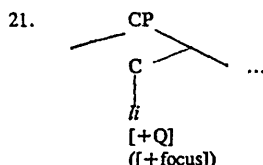
In (9) to (11), there is no focus, whereas in (12) to (15) there is. We will argue that in (9) to (11), there is no focus since the Verb incorporates in C, whereas in (12) to (15), the focus features of C are checked by the NP in Spec CP. Assuming *li* is a question marker and that it occupies C, one might argue (and we will) that the focus features in C are separate from the question features: if C merely has question features, the Verb moves; if it has focus in addition to question features, an NP must check the focus. The same is true in embedded sentences as (16) and (17) show:

16. Ne znaja filma Marija li gleda,  
'not I know whether it is Marija who is watching the film'.
17. Popitah gi dade li Ivan knjigata,  
'I asked them whether Ivan actually had given the book'.

Applying Rochemont & Culicover (1990), the question in (18) can be rephrased as in (19) but not as in (20). This shows that the pre-*li* position is indeed a focus position:

18. Koj gleda filma?  
'Who is watching the film'?
19. Marija li gleda filma?  
'Is it Marija who is watching the film'?
20. %Marija filma li gleda?  
'As for Marija, is it the film that she is watching'?

A structure for question and focus structures using CPs is (21):



In (9), the verb *gleda* 'watches' incorporates in C, checks the question features and the result is a regular question. In (10) and (11), *Marija* and *filma* move to a position preceding V and the result is topicalization of these elements, but no focus. In Rudin, the NP receives focus by Spec-head agreement, but native speakers maintain that no focus is involved. In (12) and (13), it is the preposed element that is questioned and this can be accounted for if the preposed element is in Spec CP and the features of C are checked by these NPs rather than by the Verb as in (9), (10) and (11). In (14) and (15), the first NP is topicalized but the second NP is questioned and focussed which fits if the NP checks the question and focus features.

Sentences (22) and (23) show that structure (21) must include at least three CPs. This brings up a difficult question as to the structure of these sentences. There are two possibilities: adjunction to CP or movement into (additional) Spec CPs. For this reason we examine other pre-*li* elements in the next section.

22. *Marija filma gleda li?*  
 'As for Marija and as for the film, is she watching it?'
23. *Marija knjigata na Dessislava dade li?*  
 'As for Marija and as for the book and as for Dessislava, did she give it to her?'

### 3.2 C as focus and multiple WH

As in other Slavic languages, Bulgarian obligatorily moves *wh*-phrases to a pre-C position, as (24) and (25) show. In (24), *kude* follows *če* and therefore, the sentence is ungrammatical; in (25), *li* follows *kude* and the sentence is grammatical:

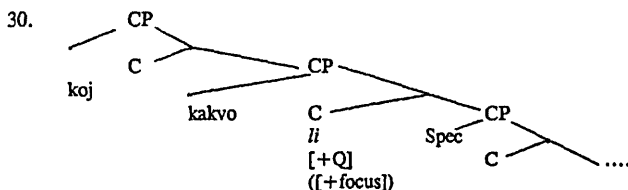
24. \**Toj izkreštja če kude e knjigata,*  
 'He shouted that where is the book'.
25. *Čudja se Ivan včera kude li beše,*  
 'I wonder Ivan yesterday where li was', i.e. I wonder where Ivan was yesterday'.

In main clauses with *li* such as (27) and (28) and subordinate clauses such as (29), however, it seems as if *wh*-elements need not move to a pre-*li* position:

26. *koj kakvo na kogo dade včera?*  
 'Who, what, to whom gave yesterday'.
27. *koj li kakvo na kogo dade,*  
 'Who li what, to whom gave'.

28. *koj kakvo li na kogo dade*  
 'Who, what li, to whom gave'.  
 29. *Ne znaja kakvo li kude stava tuk.*  
 'not I know what li where happens here'.

To account for these sentences, there are two possibilities, as mentioned before: (a) there is only one CP and *li* is its head, and (b) there are multiple CPs with one principle C. Under option (a), there is one Specifier position which can be filled by an NP or *wh*-element and *li* is an affix that moves up in, for instance, (27) and attaches to *koj*. The checking of the question and focus features occurs after affixation. The problems with this account are that checking through affixation lacks independent motivation and that *li* adjoins to the right of elements whereas Verbs adjoin to the left of *li*. In Minimalism, movement occurs only if there is a need to do so and in (27), one would be hard-pressed to find a motivation. Hence, we will argue possibility (a) is theoretically problematic. Possibility (b) involves a structure as in (30) where one C is the principle one:



The problem with (30) and possibility (b) is how to restrict Cs, but we will assume this can be done and propose that *li* checks Q-features or Q and focus-features. In the first case, a Verb or *wh*-element checks the features; in the second case, an NP does.

The problem of multiple CPs is not unique to Bulgarian since languages with multiple *wh*-phrases have the same problem. For instance, (31) needs to accommodate three *wh*-elements:

31. What did Emma talk about to whom when.

Chomsky (1992) provides no solution. He argues that C contains a question operator-feature and that this feature must be checked by an Operator. One could argue that the *wh*-elements *in-situ* in (31) are not Operators but are licensed some other way<sup>4</sup>.

Our analysis is similar to Rudin's (1993/4) and King's (1994) but differs in a number of respects. Rudin argues, like us, that apart from the interrogative function *li* is a focus marker and must discharge its focus feature at s-structure. *Li* is assumed to be generated in C and a

<sup>4</sup> Since, in English, *wh-in-situ* typically appear in focus positions, i.e. in embedded positions, one might argue that the focus-feature 'licenses' the Operator. This would not work for Bulgarian, however.

constituent is required to move to C or to Spec CP to receive the focus feature and thus to license *li*. We disagree with this aspect, which is similar in King's account, because in sentences such as (9), focus is not relevant.

King (1994) concentrates on *li* in Russian, whereas we are interested in the CP system. King argues that in Russian *li* is a clitic complementizer assigning a focus, but this cannot always be true at least for Bulgarian, given the interpretation of (9) which is one in which focus is not relevant. In Russian, one of two elements can appear before *li*, either the Verb or an XP. In the former case, the Verb adjoins to *li* in C and the entire sentence is questioned; in the latter case, an XP is moved into Spec CP before *li* and the focus is on the preposed element in Spec CP as in:

32. *vodku li ona kupila*  
'Did she buy VODKA'. (King 1994: 96)

Additional topicalized elements in Russian appear to the right of *li*, i.e. adjoined to IP as in:

33. *Ivan li ètu gazetu čital*  
'Did IVAN read this paper'.

King argues, *li* is in C; and only when an NP is directly in front of *li* is this element focussed. King proposes that the clitic contains an inherent focus feature which it can assign to an XP or to X. However, as mentioned above, the X is never focussed and therefore we prefer an analysis that separates question and focus features. For Bulgarian, King assumes an analysis similar to Russian, except that certain elements can precede the Verb and are "adjoined to CP in topic position" (p.113). We argue that there is no adjunction, just a stacking of CPs where each of the CPs can contain *li*. The parametric between Russian and Bulgarian is that there is no stacking in the former.

In Serbo Croatian, the Verb incorporates/adjoins to *li*, but NPs cannot. This indicates that *li* only contains Question features:

34. *Gleda li film Marija?*  
'Is Marija watching the film?'
35. *Je li Marija gleda film?*  
'Is it Marija who is watching the film?'

Thus, in Serbo Croatian and Russian, only one CP occurs, whereas in Bulgarian, more do. *Li* in Russian and Bulgarian is a holder for question and or focus features whereas in Serbo Croatian, focus features are not connected with *li*.

In section 3, we present data on Bulgarian focus. *Li* contains focus and/or Question features that must be checked. We have argued for multiple CPs since *li* may float.

#### 4 Nali/dali

In this section, we briefly discuss forms related to *li*. *Li* is not placed sentence-initially but *nali* and *dali* are, as (36) and (37) show. One might also say *na* and *da* fulfill the same role as the Verb in, for instance (9), namely they check some of *li*'s features through adjunction to C. Neither of the two can be preceded by a V as (38) shows but they can be preceded by an NP as in (39). This confirms our earlier analysis that Verbs, like *da* and *na*, incorporate into C to check the question features. When an NP precedes *nali* or *dali* in (39), this NP, like the NPs in (10) and (11), is not focussed. The gloss in (39) indicates that:

36. Nali razbra kakvo kazah?  
 Didn't you understand what I said'.  
 37. Dali toj šte dojde?  
 'If he will come, i.e. will he come'.  
 38. \*Razbra nali kakvo kazah?  
 39. Toj dali šte dojde?  
 'Will he come', but not: 'Is it he that will come'.

The same happens in subordinate sentences. For instance, in (16), *toj* 'he' is in focus:

40. Ne znaja dali toj ste dojde,  
 'I don't know whether he will come'.  
 41. ne znaja toj dali ste dojde.

Both *dali* and *nali* derive diachronically from modal and negation particles respectively, merged with the question particle *li*. We do not elaborate on the functional categories these particles may originate in.

#### 4 Conclusion

Bulgarian focussed and topicalized elements can be accommodated in a structure with stacked, i.e. multiple, CPs. The focus features in C (and question features) can be checked by an NP in Spec CP. When C merely contains question features, it is the Verb that checks these (as in other languages). The topicalized elements always appear to the left of the structurally focussed NP. There is evidence that the focus features can be in any of the Cs, as in (26) to (28) and this indicates that the Functional projections are CP rather than FocusPs. Universal Grammar provides building blocks such as CPs and focus features but languages vary as to how they are used. The problem is restraining too many independent Cs.

The notion of multiple CPs claimed by us to account for Bulgarian focus has been argued for before for different languages. Authier (1992), for instance, has argued for embedded topicalization in English; de Haan and Weerman (1986) for Frisian; Platzack (1986) and Kosmeijer (1993) and Hoekstra (1993) have for other Germanic languages. In these languages the possibility of multiple CPs has been linked to some properties of the C (e.g. Verb-second in embedded sentences occurs after verbs like *say*, see also van Gelderen 1993). In Bulgarian, multiple CPs occur with fewer constraints.

Parametric differences between Russian, Bulgarian and Serbo Croatian involve the status of *li* and the presence of CPs. In Serbo Croatian, *li* is only used as question particle whereas in Russian, it can also be used as a focus marker. The latter two languages only have one CP, however.

## References

- Authier, J.-M. 1992. "Iterated CPs and Embedded Topicalization", *Linguistic Inquiry* 23.2: 329-36.
- Chomsky, N. 1992. "A Minimalist Program for Linguistic Theory", *MIT Occasional Papers in Linguistics* 1.
- Chomsky, N. 1994. "Bare Phrase Structure", *MIT Occasional Papers in Linguistics* 5.
- Cinque, G. 1993. "A Null Theory of Phrase and Compound Stress", *Linguistic Inquiry* 24: 239-97.
- Gelderen, E. van 1993. *The Rise of Functional Categories*. Amsterdam: John Benjamins.
- Haan, G. de & F. Weerman 1986. "Finiteness and Verb Fronting in Frisian", in H. Haider & M. Prinzholm, *Verb Second Phenomena in Germanic Languages*, Dordrecht: Foris: 77-110.
- Hoekstra, E. 1993. "Dialectal Variation inside CP as Parametric Variation", *Linguistische Berichte* 5.
- Horvath, J. 1986. *Focus in the Theory of Grammar and the Syntax of Hungarian*. Dordrecht: Foris.
- King, T. Holloway 1994. "Focus in Russian Yes-No Questions", *Journal of Slavic Linguistics* 2.1: 92-120.
- Kosmeijer, W. 1993. *Barriers and Licensing*. Groningen PhD.
- Platzack, C. 1986. "Comp, Infl and Germanic Word Order", in L. Hellan & K. Koch Christensen (eds), *Topics in Scandinavian Syntax*. Dordrecht: Reidel.
- Rivero, M.-L. 1993. "Bulgarian and Serbo-Croatian Yes-No Questions: V<sup>o</sup>-Raising to -LI versus -LI Hopping", *Linguistic Inquiry* 24.3: 567-75.
- Rochement, M. & P. Culicover 1990. *English Focus Constructions and the Theory of Grammar*. Cambridge: CUP.
- Rudin, C. 1988. "On Multiple Questions and Multiple WH Fronting", *Natural Language and Linguistic Theory* 6: 445-501.
- Rudin, C. 1993/4. "On Focus Position and Focus Marking in Bulgarian Questions", ms.
- Tuller, L. 1992. "The Syntax of Postverbal Focus Constructions in Chadic", *NLLT* 10: 303-334.





This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).